CONFIDENTIAL

DD/P 4-5276

13 OCT 1964   DDI 3967-64

MEMORANDUM FOR: Mr. Robert I. Bouck
        Special Agent in Charge,
        Protective Research Section
        Secret Service

SUBJECT   : Automatic Data Processing of Biographic
        Intelligence

REFERENCES : A. CSCI-3/779,271
        B. CSCI-3/781,077
        C. Letter from Mr. James J. Rowley to
          Lt. General Marshall S. Carter dated 13 May 1964

   The Committee on Documentation (CODIB) of the U.S. Intelligence Board has established a task team to examine the application of EDP techniques to biographic intelligence processing. Attached are the minutes of the first meeting of the team and a draft of its terms of reference.

   In light of the Warren Commission recommendation that the Secret Service should attempt to develop a system compatible with those of the agencies from which most of its data will come, you are invited to participate in subsequent meetings of the CODIB task team. Although your biographic collection may not yet involve millions of cards, information obtained by the team on methods and techniques being developed in the community should be of interest.

   The workings of the task team will necessarily stretch out over a considerable period of time. Meanwhile, we would be more than happy to follow up on your briefing of 8 May with more detailed explanations of the Clandestine Services automated document retrieval and name indexing systems with you and your technical staff.

            (Signed) Richard Helms

             Richard Helms
           Deputy Director for Plans

Encl.

CONFIDENTIAL

Group 1
excluded from automatic
downgrading and
declassification

☐ UNCLASSIFIED ☐ INTERNAL USE ONLY ☐ CONFIDENTIAL ☐ SECRET

ROUTING AND RECORD SHEET

SUBJECT: (Optionol)

Automatic Data Processing of Biographic Intelligence

FROM: DDP/SG 3-C-19		EXTENSION	NO.
			DATE / 3 October 1964

STATINTL

TO: (Officer designation, room number, and building)	DATE		OFFICER'S INITIALS	COMMENTS (Number eoch comment to show from whom to whom. Drow o line across column ofter eoch comment.)
	RECEIVED	FORWARDED		
1. 1 7 E 32 *ADDI/M*				25X1A File CODIB TT/V
2.				
3. *CODIB Secretary* *2 E 60*				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
11.				
12.				
13.				
14.				
15.				

25X1A
25X1A
25X1A

FORM 3-62 **610** USE PREVIOUS EDITIONS   ☐ SECRET   ☐ CONFIDENTIAL   ☐ INTERNAL USE ONLY   ☐ UNCLASSIFIED

S-E-C-R-E-T

J Biographics)

10 January 1962

PROPOSAL FOR A BIOGRAPHIC INFORMATION NETWORK

INTRODUCTION

It is now possible to construct a central facility to provide quick access to information on foreign personalities in the intelligence community. If work were to commence in all Agencies by the mid 1962, it could be in partial operation by 1965. Pooling of subst........ information is not involved. Compartmentation and need-to-know security would be assured by housing indexes in electrically linked but separately operated computers, each programed by personnel of the owning Agency. The only new permanent facility would be a switching and message center to operate the communications network, and a medium-sized computer. containing alphabetical and phonetic look-up tables. Intelligence officers and systems engineers representing their parent agencies would form a committee to keep programs and equipment compatible, and the look-up tables up-to-date.

The proposal is presented (in this issue of the STUDIES IN INTELLIGENCE, as well as separately to the Committee on Documentation of the USIB) in order to draw attention to our opportunity. A long lead time and much preparatory effort are needed to exploit it to the community's benefit. In particular, there is an immediate requirement to

S-E-C-R-E-T

S-E-C-R-E-T

-2-

capture the typing of biographic index cards or forms in machine language
--such as punched cards or paper tape--in those Agencies which consider
themselves potential participants in the system.

The scope of this participation cannot be outlined in this
paper with any precision. The writer's perspective derives from the
counterintelligence field, but he sees no reasons why the principles of
index organization and information handling which are outlined below
might not equally apply to the positive biographic intelligence area.
In practice, the dividing line between the two is elusive; we have long
since given up the quest for a durable definition of the term "derogatory
information."

The great bulk of name searching activity in the intelligence
community--thousands of requests daily--is initiated to determine whether,
in the records of any one of a number of Government agencies, there is
information of a nature to preclude employment, the receipt of a grant,
the issuance of an immigration or a visitor's visa, entry to the U.S.,
an invitation to a conference, etc. Only in a fraction of name searches
is such information discovered; most people are quite all right.

Using modern communications and data processing tools, the com-
munity can rid itself of at least that mountain of paper work which now
piles up merely to determine that no pertinent information is known. It
should be feasible routinely to receive negative answers in an hour or
less. Much Government business would then promptly proceed which is now

S-E-C-R-E-T

held up for weeks. Many people the world over who want to serve or visit the United States would be relieved, impressed, and grateful.

This paper will demonstrate, furthermore, that the quality of name tracing by present methods is not as high as may be generally believed. Due to inevitable shortcomings in manpower and qualifications, and the inroads of human error, there is a fair amount of information in the files which cannot be found. Techniques are becoming available which can produce more comprehensive and reliable replies to inquiries about people.

These techniques, once developed and applied to a single large biographic index, can be simply and inexpensively adapted--without violating the rules of security compartmentation--to the indexes of other organizations.

25X1A

Next 12 Page(s) In Document Exempt

S-E-C-R-E-T

-17-

FUNCTIONS OF THE NETWORK CENTER

A Pool of Name Tables

The term "network" is used to make it clear that the system proposed in this paper does nothing more than to relay inquiries and answers within the limits of the computer programs developed by the participating Agencies. The Center of the network does not deal with substantive information; it does no professional work other than to maintain dictionary tables; it does not even see substantive information developed by a participating Agency in response to a request.

The central idea is to capture all surnames occurring in the indexes of the participating Agencies as a byproduct of typing and machine processing, and to pool them in central computer tables. They are thus divorced from all other information held by the Agencies concerned. The tables are then to perform two functions: to take care of the name variant problem centrally in the same manner as outlined above; and to direct searches to those Agencies which, according to the central tables, have filed information on someone by either the specific surname filed at the Center, or by any of its grouped variants. Participants could, of course, direct their inquiries to one or several agencies specifically.

Given name tables and other dictionaries would also be stored in the Center's computer. Their purpose is to obtain a mechanical translation from a given name (or other search element) to a number, or a code,

S-E-C-R-E-T

which stands for all its versions in each of the participating Agencies' computer indexes.

Creation of Central Tables

The grouping of names in the counterintelligence index of CIA is apt to reach a plateau sometime in 1963, after which relatively few name spellings which have not occurred before will have to be dealt with. From then on, the Agency will be in a position to supply surname and given name magnetic tapes to the central computer in both alphabetic and group number order. (This paper is an outline of possibilities, not an official offer to this effect.)

Surnames processed in the computers of other Center customers can next be matched against those of CIA. Those on which a match occurs will be tagged with a symbol of the participating agency. If there is no match, a punched card is produced. Linguists use it to assign the appropriate group number, thus rounding out the Center's tables.

In return, the Center will furnish the contributing agency computer materials--tapes and programs--for the organization and storage of its own index entries in groups corresponding to the surname tables at the Center.
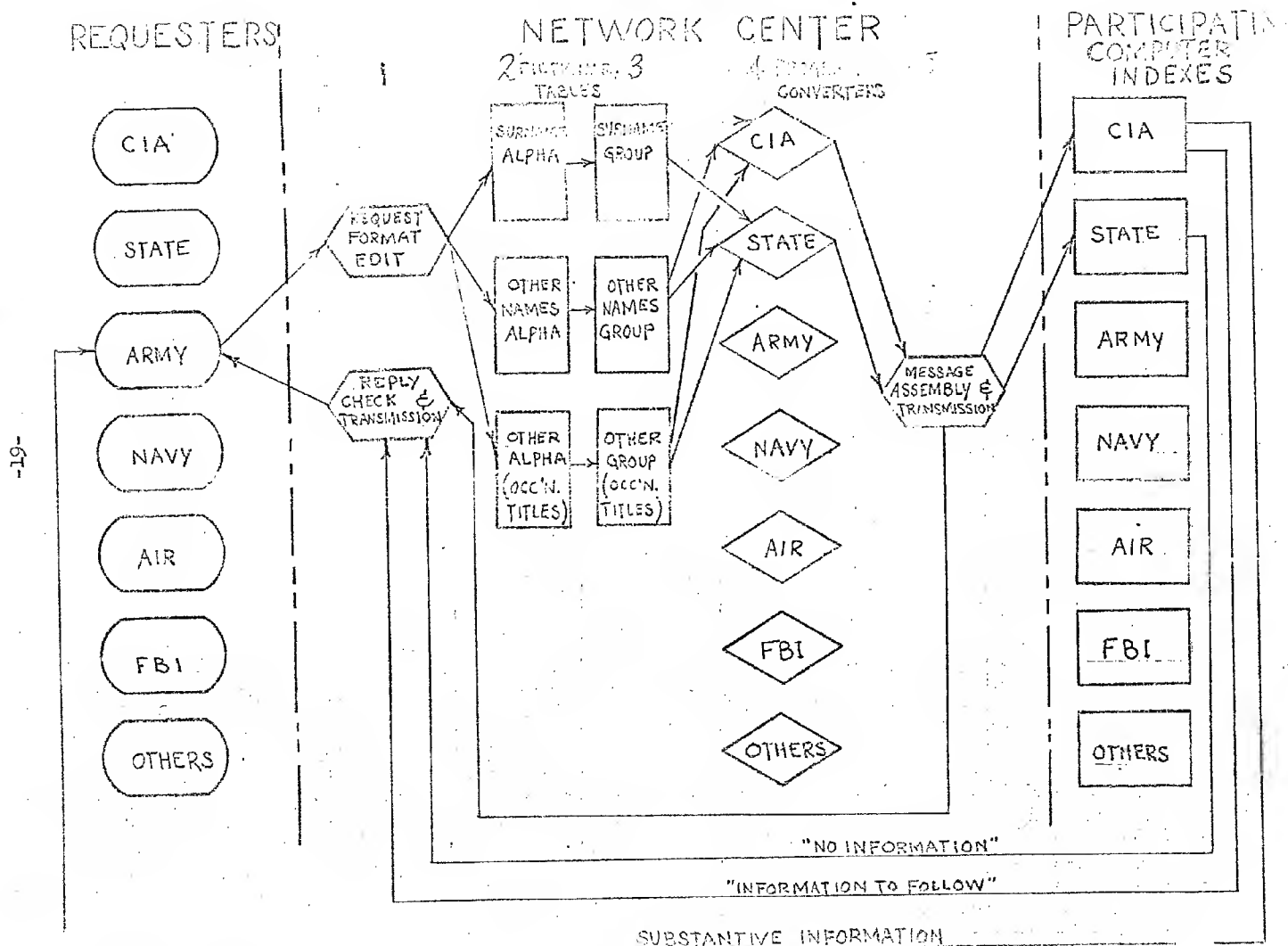
A Search through the Center

The chart on the next page illustrates the Center's modus operandi. We may anticipate these steps:

a. Using a predetermined common format, a member agency teletypes

SECRET

REQUESTERS | NETWORK CENTER | PARTICIPATING COMPUTER INDEXES

CIA

STATE

ARMY

NAVY

AIR

FBI

OTHERS

REQUEST FORMAT EDIT

REPLY CHECK & TRANSMISSION

TABLES

SURNAME ALPHA → SURNAME GROUP

OTHER NAMES ALPHA → OTHER NAMES GROUP

OTHER ALPHA (OCC'N. TITLES) → OTHER GROUP (OCC'N. TITLES)

CONVERTERS

CIA

STATE

ARMY

NAVY

AIR

FBI

OTHERS

MESSAGE ASSEMBLY & TRANSMISSION

CIA

STATE

ARMY

NAVY

AIR

FBI

OTHERS

"NO INFORMATION"

"INFORMATION TO FOLLOW"

SUBSTANTIVE INFORMATION

SECRET

-19-

its search request to the Center. It contains the usual elements: Name, date and place of birth, sex, citizenship, residence, occupation, etc. It includes information for human use once computers have produced possibly relevant references, such as purpose of the search, its intended depth, additional information about occupation, geography, events, etc. which will assist the analyst.

b. When the message is in, the Center's computer first assigns a serial number which accompanies the processing of all request elements until an answer is returned.

c. Then it separates those request elements which require matching against the Center's tables. The match on a surname yields not only a group number, but also a roster of Agencies in whose indexes information may be found.

d. Next, the machine loads the complete request, together with the group numbers it looked up in the tables, into the "program converters" of those member Agencies whose indexes are to be checked. These program converters are sections of the computer in which the common request format is transposed into that used by a participating Agency for its own index machine records. There is thus no need for total systems uniformity among the members.

e. Message assembly, transmission to the members' computing installations, and bookkeeping are the remaining steps in processing the search to the members' mechanized indexes.

f. An answer will be expected from each of the indexes addressed;
a log in the "Reply Check" section of the Center's computer will call a
Center employee's attention to delays beyond a reasonable time of, say, an
hour. When the returns are in, the computer composes a message to the
requester, and transmits it.

g. These answers are of two types only:

(1) No pertinent information available, or

(2) Possibly pertinent files are being reviewed; further
information will come through normal, direct channels.

h. There are two possible sources for the completely negative
answer: One is the member Agency's index computer itself, drawing a
blank on the search criteria. The response in this case is automatic.
If it yields possibly pertinent references, an analyst should rapidly
scan these merely to determine whether it might be useful to consult
documents or files. If not, the search ends at this point. A simple
instruction from the analyst to the computer will then advise the
requesting Agency whether to expect more information or not.

Is It Worth Doing?

Mountainous computers have labored, and brought forth one of
two mice--either a negative, or a promise of something to come, which
may once more turn out to be negative. Otherwise put, all we are suggest-
ing is an immediate show of ignorance, plus an expedited review of perti-
nent information leading to the same sort of positive reply now produced

by analysts in all Agencies.

The economy of computer systems is often a touch-and-go matter. With no information from outside CIA available, we have, of course, no proper basis on which to make a calculation. However, let's play with some numbers anyway.

We do over one thousand name traces daily in CIA headquarters alone, most of which are run as well in several other Agencies. The community surely makes 15,000 trips a day to its name indexes and files, not counting their look-ups in overseas repositories. (Our plans for the future include tying these into the headquarters computer facility through electrical communications.)

It appears to take us about an hour in our present manual system to do the average name trace. The community's aggregate manpower commitment to name tracing is thus apt to be around 15,000 man/hours daily. However, the 10 or 15% of all traces which require file review or memo writing probably take up half of this time.

The bull's eye we are shooting at is thus on the order of 7,500 man/hours a day, or the labor of around 1,000 people, who cost the Government some $10,000,000 annually in salaries and overhead. The aggregate computer budget should remain below this figure. It should be worth doing in the name of economy, as well as long-term quality and service gains.

How To Get Started

We have listed neither all the benefits, nor all the costs of

S-E-C-R-E-T

-23-

such a system. Among the latter, conversion to machine language of exist-
ing indexes is a prominent factor, so much so that it will have to be
spread over a long period of time.

We began typing our index cards with a machine language byproduct
in the fall of 1957. By early 1964, when our index computers should begin
to operate, we will probably have half--the recent, most useful half--of
our index entries in machine language. Then begins the drive on the rest.

What if we had not begun in 1957? We would be overawed by the
conversion problem, instead of merely impressed and annoyed. And the pros-
pect would appear less pleasant every day. Optimists might point to the
development of mechanical print reading techniques; we have little faith
in their application to a collection of heterogeneous 3x5 index cards.
Every index or reference card now produced on a standard typewriter by a
USIB agency which later decides to use a computer--whether in the proposed
network or by itself--will probably have to be retyped. To make a begin-
ning, therefore, index and reference typists need to be re-equipped with
keypunches or paper tape-producing typewriters.

CONCLUSION

Those members of the intelligence community which have taken
this step are in a position to link their biographic indexes in a network.
Consider the trends: Information volumes are going up, as are the number
of trace requests. Clerical manpower needs and labor costs are going up

S-E-C-R-E-T

S-E-C-R-E-T

-24-

correspondingly. On the other hand, computer flexibility is improving,
and computer costs--per unit of work done and information stored--are
coming down. And as we read current history, we see no visible trend
toward a lessening of pressure on the security of the United States here
and abroad which might reduce the Government's need to use its intelligence
and counterintelligence tools as effectively as possible.